

A REPORT BY THE WORKSHOP

2019

MAY

D I G I T A L
T H R E A T S
T O
D E M O C R A C Y

The
Workshop

The
Law
Foundation
NEW ZEALAND

Luminate

D I G I T A L T H R E A T S T O D E M O C R A C Y

This report is part of the Digital Threats to
Democracy research project.

To see the rest of the reports and the
overall findings go to digitaldemocracy.nz

ISBN: 978-0-473-48026-4

CONTENTS

ACKNOWLEDGEMENTS	04
INTRODUCTION	05
PURPOSE, DEFINITIONS AND METHOD	08
Research purpose	08
Definitions	08
Methodology	09
Overview	11
THE OPPORTUNITIES OF DIGITAL MEDIA	14
Democratization of information publishing	15
Broadening of the public sphere	15
Increasing equality of access to and participation within political processes	16
Increasing engagement in political processes	16
Increasing transparency and accountability from government	16
Promotion of democratic values	16
THE THREATS	17
Increasing power of private platforms	18
Foreign government interference in democratic processes	18
Surveillance and data protection	19
Fake news, misinformation and disinformation	19
Filter bubbles and echo chambers	20
Hate speech and trolling	20
Distrust/dissatisfaction with democracy	21
THE SOLUTIONS	22
A Hierarchy for Solutions and Interventions	22
Change society wide structural & systems issues	25
Create supportive environments & contexts - making the default inclusive and safe	26
Create long lasting protections for people or intervene to protect them	28
Build understanding and change individual behaviours	29
WHAT ROLE FOR NZ?	30
New Zealand as follower	31
New Zealand as leader	31
New Zealand as niche influencer	31
CONCLUSIONS	32
Inclusive and transparent processes are critical	33
More research needed	33
Evidence-led and principled approach to urgent policy development	33
Apply human rights principles	34
Agile and responsive approach to policy	34
Urgent areas for action	35
SUMMARY OF KEY FINDINGS FROM EACH PART OF THE RESEARCH	36
Key findings from the survey	36
Key findings from the literature review	37
Key findings from the interviews	38

ACKNOWLEDGEMENTS

This research was made possible by support from our major funder, the New Zealand Law Foundation's Information Law & Policy Project, with additional research funding from the Luminate Group.

The New Zealand Law Foundation - Te Manatū a Ture o Aotearoa – provides grants for legal research, public education on legal matters and legal training. The **Information Law & Policy Project (ILAPP) funds** projects that will better prepare New Zealand for the challenges of the information age. The project is intended to support the growth, understanding and resilience of New Zealand and prepare the country for future digital competence. It will support and feed into work the public and private sector is undertaking, but will remain independent.

The Luminate Group is a global philanthropic organisation with the goal of empowering people and institutions to work together to build just and fair societies. They fund and support projects that will help people participate in and shape the issues affecting their lives, and make those in power more transparent, responsive, and accountable. Their focus is on civic empowerment, data & digital Rights, financial transparency, and independent media.

The research team on this project was lead by Marianne Elliott (The Workshop) and included Dr Jess Berentson-Shaw (The Workshop), Dr Kathleen Kuehn (Victoria University of Wellington), Dr Leon Salter (Massey University) and Ella Brownlie (The Workshop). Project management was provided by Jay Brooker (The Workshop). The quantitative survey was conducted by UMR Market Research.

We particularly want to thank all the participants in the interviews for this research. They volunteered their time on this project, brought both goodwill and clear thinking to the unreasonably broad scope of the research, and responded rapidly to our requests for feedback. This kind of generosity and rigour across the sector gives us hope that with more investment in research and a more joined-up government approach to the subject, New Zealand will be able to contribute substantially to an urgently needed global collaboration on solutions.

INTRODUCTION

In February this year, as I pored over the findings of our literature review and read through hundreds of pages of interview transcripts, I wrote that an adequate response to the problem of online hate, harassment and abuse was possible. It would require a recalibration of our policy approach, some international diplomacy and cooperation, and a sufficiently diverse group of decision-makers at the helm.

I believed then that all of that was within the capacity of the New Zealand government, and that there was “likely to be a leadership role for our country in global efforts to combat online abuse and, as Sir Tim Berners-Lee has put it, ‘fight for the web’.”

This belief has proven to be founded, although under circumstances none of us ever wanted to witness. As we completed this research, it was announced that New Zealand’s Prime Minister, Jacinda Ardern, would meet the French President, Emmanuel Macron in Paris to “bring together countries and tech companies in an attempt to stop social media being used to promote terrorism.” The meeting will invite world leaders and tech company CEOs to sign a pledge called the ‘Christchurch Call’.

In many ways, the devastation of the Christchurch mosque massacres has proven to be a turning point for New Zealand on this, and other policy issues. We now know that our small size and relative remoteness do not render us immune to the terrible harm that can be done by a person motivated by hatred, inspired by the internet and armed with a semi-automatic weapon.

In the wake of the March 15 attacks, in response not only to the unthinkable cruel and manipulative use the terrorist made of the internet in the course of the attack but also to the many ways in which online spaces have allowed hatred to grow and spread, many people - including our Prime Minister - called for greater accountability and care from the big digital platform companies, including Facebook and Google.

It’s a call some of us have been urging our government to make for some time now, and many feel it is long overdue. But here we are now, and this is a crucial moment in the history of the relationship between citizens, governments around the world and a handful of people who not only control a significant portion of the means by which we all communicate and the distribution of news and information to vast percentages of the world’s population, but also hold huge quantities of personal data about us all.

The question is no longer whether something needs to change. The question has become: what precisely needs to change? And even more importantly: what can be done? What evidence do we have as to the interventions and solutions that might mitigate against the biggest threats posed to our democracy by digital media, without losing the best of the opportunities that the internet offers? Those are the questions we set about answering with this research.

We are far from the first people to tackle these questions, as our literature review reveals. Researchers, academics, journalists and former employees of the big tech companies have been studying and writing about the impact of digital media on democracy in increasing numbers over recent years.

In his book, *The People vs Tech* Jamie Bartlett predicted: “In the coming few years, either tech will destroy democracy and the social order as we know it, or politics will stamp its authority over the digital world.” In his view, “technology is currently winning this battle, crushing a diminished and enfeebled opponent.”

Similarly, in *How Democracy Ends*, David Runciman assessed the comparative strengths of the tech giants versus governments, in a ‘Leviathan vs Leviathan’ showdown for the future of democracy. Although he gave governments more of a shot than Bartlett had, he concluded that while “Facebook will not take down the Leviathan in mortal combat ... it could weaken the forces that keep modern democracy intact.”

But neither Runciman nor Bartlett, nor any of the analysis I’ve read over the past year, predicted the situation we are now in. None imagined a Prime Minister with a global reputation for compassion, armed with moral courage, clarity and the support of an outraged nation.

Has Jacinda Ardern become the global leader capable of taming the tech giants? There are good reasons to hope so, and even more reasons to ensure that this rare opportunity is neither wasted nor lost.

NEED FOR A COHESIVE, EVIDENCE-BASED APPROACH TO POLICY

One of the challenges of rapidly developing policies on digital media in response to a situation like the Christchurch attacks is that this entire area of policy has been relatively neglected until recently. As one interviewee in this research said, we need a better system for making policy on these issues before we can be any kind of global leader. “Smart people just basically giving their opinion with no real information behind it,” won’t be good enough to develop the kind of solutions demanded by this particular set of problems and, they say “that’s how we’ve made our policy in this space, generally.”

Until very recently, there was no centralised or coordinated government process for developing policies and strategies in response to the challenges posed by digital media. Responsibility fell to a wide range of different agencies and teams, and policy development was consequently, inevitably, fragmented. In the process of doing this project, we found it difficult to establish who in government, if anyone, had a broad view over the full range of issues raised in our research. Recently, new efforts at coordination have begun to appear, with some degree of overarching responsibility, although not necessarily with the resources needed to develop policy across such a wide-ranging and rapidly changing area.

In the past, according to one interviewee, New Zealand has either simply adopted the policy approach taken in another jurisdiction “or we have a relatively flimsy policy discussion which isn’t founded in evidence.” In order to build our capacity as a country to understand and deal with these issues, they argue, we need more of an evidence base. “Before we can be leaders in any sense, we need to be equipped to have a solid base for developing policy ourselves.”

What our research shows is that it is critical that the Prime Minister and her advisors look beyond immediate concerns about extremism and content moderation, and ensure that our government's efforts in this moment take into account the wider structural issues that created the conditions in which a live video of an act of such violence could be shared and viewed so widely.

Those wider structural issues include in particular the impact of **platform monopolies**, in which a handful of people have the power to determine the social interactions and access to information of millions of people, **algorithmic opacity**, in which algorithms have ever-increasing influence over what we hear and see without appropriate transparency or accountability, and the **attention economy**, which gives priority to content that grabs attention, without sufficient regard to potential harm.

Our intention is that this research will contribute to a wider consideration of the issues arising from digital media's impact on democracy, and to the development of a body of evidence which supports this critical work.

Marianne Elliott

Lead Researcher, Digital Media and Democracy

Co-Director, The Workshop

PURPOSE, DEFINITIONS AND METHOD

RESEARCH PURPOSE

Digital media has been heralded as inherently democratising. People have direct access to each other and to their elected representatives, across geographical and cultural boundaries. But increasingly it is also seen as a space in which democracy may be simultaneously undermined.

As digital technology increasingly permeates society, there is good reason to pay attention to the institutions, policies, and practices that surround this technology and present both opportunities and threats to democracy. This is especially true for government, as those people who represent the interests of all citizens, but it is also true for everyone with an interest in the future health of our democracy.

The purpose of this research was to explore the opportunities, risks and threats posed to New Zealand's democracy by digital media, in order to scope future research into the policy solutions available to New Zealand to maximise the opportunities, and to meet and mitigate the threats.

DEFINITIONS

In order to assess the impact of something like digital media on democracy, you need a definition of democracy. We used a definition of democracy adapted from the framework developed by the Economist Intelligence Unit for their Democratic Index report, and the definition used by Jamie Bartlett in his book 'The People vs Tech'.

The five features of democracy in our definition are:

Electoral process and pluralism: including whether elections are free, fair and trusted.

Active citizens: alert, informed citizens who are capable of making important moral judgements, including measures of equity and diversity in representation.

Shared democratic culture: enough societal consensus, cohesion and willingness to compromise for a stable, functioning democracy. In New Zealand, this includes

compliance with te Tiriti o Waitangi, on which our democratic culture is founded.

Civil liberties and competitive economy: a functioning competitive economy and civil society, including protection of human rights and free, independent media.

Trust in authority: a trustworthy government, parliament and judiciary and elected representatives accountable to the people.

Defining democracy is complex. Defining digital media is almost as difficult. Digital media technically includes all digitised content that can be transmitted over the internet or computer networks. This could include text, audio, video, and images. So content from print or broadcast media outlets can fall into this category when it is presented on a website or blog.

The focus of this research was on social media, online forms of communication that people use to share and exchange information with interested audiences, and within that, a specific focus on the major digital platforms, including Facebook, YouTube and Twitter. However, interviewees also talked about the impact of other forms of digital media on democracy, including blogs, online forums and digital forms of traditional media.

METHODOLOGY

This research project was made up of three separate but related strands: expert interviews, a literature review, and a quantitative survey.

EXPERT INTERVIEWS

A series of in-depth interviews were conducted with experts and stakeholders to explore the scope of this issue in more detail, prioritise various aspects of the problem for future research and identify key potential collaborators for further research.

Thirty-five in-depth interviews were undertaken with a selection of experts, stakeholders and users drawn from the following sectors.

- Political
- Policy and official
- Māori-led organisations
- Civil society
- Industry/sector organisations
- Academics, researchers and experts
- International experts.

The interviews were recorded and transcribed and analysed using a hybrid of content and grounded analysis in which some broad themes were used as a starting framework for the analysis, but amended and altered based on the themes that emerged from the data as the analysis progressed.

The full report on the interviews can be accessed [here](#)

LITERATURE REVIEW

The literature was conducted in two parts, one looking at the nature of the opportunities and threats to democracy from digital media, and the other looking at the evidence as to effective solutions and responses.

Part one: Opportunities and Threats

In this narrative literature review, we sought to describe, from the most recent literature (searches were limited to research published in the last eight years and most are within five), what the nature of the opportunities and threats are to democracy from developments in digital media.

We asked two research questions:

1. What are the specific opportunities digital media presents for improving democratic participation?
2. What are the current threats/barriers that are in place to prevent achieving those opportunities?
3. A non-systematic narrative review was chosen with a view to summarising the themes that have been covered in terms of opportunities and problems (risks and threats). Searches were limited to research published in the last eight years (most are within five).
4. In total, 110 documents were reviewed including journal articles, reports and book chapters.

Part Two: Solutions

Following on from the review of the literature identifying the opportunities and threats that digital media pose to an inclusive and participatory democracy (Part One), we undertook a review to identify tested and workable solutions to realising the potential of digital media and/or overcoming current threats.

A non-systematic narrative review was chosen with a view to summarising the evidence. Searches were limited to research published in the last eight years (most are within five). It was not an exhaustive review, but in general we found a dearth of empirically tested solutions. This is not surprising given the relatively slow response of government and other public institutions (from where such research would most logically be situated and or funded) to the threats from digital media.

The review is presented in three parts: 1) the empirical evidence on workable solutions to threats to democracy from digital media, 2) a summary of recommendations found in the literature and 3) a brief discussion of some activities identified in New Zealand

The full literature reviews can be accessed [here](#).

QUANTITATIVE SURVEY

The third component of this research was an online survey among a nationally representative sample of New Zealanders aged 18 years and over. The survey was designed to elicit the views and experiences of people using social media and digital platforms relevant to democracy (e.g. participating in debates about issues of public policy on social media.)

1,000 people were surveyed, weighted to accurately reflect the New Zealand population in relation to region, age, gender and ethnicity. Fieldwork was carried out from the 27th of September to 2nd of October 2018.

The full report on the survey can be accessed [here](#).

OVERVIEW

Over recent years a growing body of international research has looked at the impact of digital media on democracy, with particular focus on the US and the UK where the role played by digital media in the election of Trump and the Brexit referendum raised significant concerns.

This project was designed to find out if we should be worried about these same issues here in New Zealand. And if so, what should we do about it? In order to answer that question we identified five key features of democracy against which we could measure the impact of digital media, for better and for worse. They are:

- Electoral process and pluralism
- Active, informed citizens
- Shared democratic culture
- Civil liberties and competitive economy
- Trust in authority

WHAT WE'VE FOUND

Critically, we found that digital media is having an impact across every one of those features of a healthy democracy.

There are indicators that digital media has had some beneficial impacts. Our quantitative research here in New Zealand indicates, for example, that people from minority groups have been able to use digital media to participate in democratic processes including accessing politicians and engaging in public debate. Whatever our response to the challenges posed to democracy by digital media, it's important we don't lose these opportunities in the process.

But the overall trend should raise serious concerns. Active citizenship is being undermined in a variety of ways. Online abuse, harassment and hate - particularly of women, people of colour, queer people, people with disabilities and people from minority religions - undermines democratic participation not only online, but offline.

Misinformation, disinformation and mal-information are undermining not only informed debate, but also public trust in all forms of information. Distraction and information overload are eroding citizens' capacity to focus on important and complex issues, and their capacity to make the 'important moral judgements' required in a functioning democracy.

Likewise, interviewees described a myriad of ways in which our shared democratic culture is being undermined by digital media - including through disinformation, polarisation, attention hijacking and radicalisation.

One of the clearest impacts of digital media on our democracy has been its impact on funding for mainstream media. While Facebook and Google Hoover up the advertising revenue that once would have been spent on print, radio and television advertising, they contribute nothing to the work of producing the kind of news and current affairs reporting that is essential to a functioning democracy. In a stunning display of hypocrisy, Facebook recently complained that their local news service was being hindered by a lack of local newspapers, many of which were forced to either shut down or significantly reduce their newsroom size after losing advertising income to Facebook.

The representative survey we carried out indicates that New Zealand's small size and relatively healthy mainstream media (relative to elsewhere and despite significant resource challenges) may help us avoid the worst effects of "filter bubbles" and "echo chambers" in digital media on some issues.

When asked about the legalisation of cannabis, New Zealanders who got their information about the issue online were able to predict relatively accurately whether the majority of New Zealanders shared their views or not. A third of those who disagreed could predict (that is a minority), most who agreed could accurately predict. This may be unique to the debate about drug reform because, for example, there had been significant media coverage of opinion polls on this issue. More research would be needed to see if this is replicated across other issues in New Zealand.

Interviewees in our qualitative research nonetheless pointed to examples where debate in New Zealand about issues like free speech, hate speech and gender identity attracted the attention of foreign actors holding strong, even extreme, views on these issues. Engagement by these foreign actors in the online public debates on issues here in New Zealand appears to some interviewees to have contributed to a polarisation, even radicalisation of views here. Interviewees also raised concerns that the ability of citizens to form free and informed opinions were being undermined not only by mis and disinformation, but by the increasing role of algorithms in predicting and curating the information each of us is exposed to.

THE NEED FOR A SYSTEMIC RESPONSE

We could continue to outline the impact digital media is having on trust in public institutions, free and fair elections, the protection of human rights and a competitive economy. More on all of that below. The key message is clear, digital media is having massive, system-wide impacts on our democracy. It affects every part of our lives and the people who run the corporations controlling the major platforms are having a determinative impact on the very structures and functions of our society. While better content moderation is clearly one of the responses we must demand of the platforms, it is not even close to being a sufficient response to the scale of the challenge.

THE THREE CORE PROBLEMS TO EMERGE FROM OUR RESEARCH

At the heart of the challenges to democracy posed by digital media are three core problems:

5. **Platform monopolies:** two or three corporations control not only our means of communication, but also the content which is distributed both of which are core aspects of our democracy, whilst the market power and global mobility of these companies make it possible for them to avoid national regulatory measures either by moving operations elsewhere or simply ignoring them;
6. **Algorithmic opacity:** algorithmic engines are using huge quantities of personal data to make ever more precise predictions about what we want to see and hear, and having ever increasing influence over what we think and do, with little transparency about how they work or accountability for their impact; and
7. **Attention economy:** the dominant business model of digital media prioritises the amplification of whatever content is best at grabbing our attention, while avoiding responsibility for the impact that content has on our collective wellbeing and our democracy. And the negative impact is brutally clear from both the literature and the world around us.

It's critical that this moment of global cooperation is used to address the wider, structural drivers of the biggest threats posed to democracy by digital media. These structural drivers include the power that a handful of privately-owned platforms wield over so many aspects of our lives, from what information we see, who we interact with, and who can access information about us. And we must do this while maintaining and building upon the many opportunities digital media simultaneously offer to tackle some of the biggest challenges facing democracy, including inequity of access and declining engagement.

In order to do that, action is needed sooner rather than later in order to:

- **Restore** a genuinely multi-stakeholder approach to internet governance, including rebalancing power through meaningful mechanisms for collective engagement by citizens/users;
- **Refresh** antitrust & competition regulation, taxation regimes and related enforcement mechanisms to align them across like-minded liberal democracies and restore competitive fairness, with a particular focus on public interest media;
- **Recommit** to publicly funded democratic infrastructure including public interest media and the creation, selection and use of online platforms that afford citizen participation and deliberation;
- **Regulate** for greater transparency and accountability from the platforms including algorithmic transparency and accountability for verifying the sources of political advertising;
- **Revisit** regulation of privacy and data protection to better protect indigenous rights to data sovereignty and redress the failures of a consent-based approach to data management; and
- **Recalibrate** policies and protections to address not only individual rights and privacy but also collective dynamics and wellbeing, and protect indigenous rights. Public agencies responsible for protecting democracy and human rights online should reflect, in their leadership and approaches, the increasing diversity of our country.

THE OPPORTUNITIES OF DIGITAL MEDIA

The potential of digital democracy lies in its ability to increase democratic participation, embrace diversity of opinion, and empower marginalised groups. We identified six clear opportunities from the literature that digital media offers. These are: the democratisation of information publishing, broadening the public sphere, increasing equality of access to and participation within political processes, increasing participation and engagement in political processes, increasing transparency and accountability from government and promotion of democratic values.

Broadly speaking these opportunities fit into two categories: First, those that enable individuals, citizens or groups, who due to their status in society have been excluded from fully participating in different aspects of the democratic process, through greater access to the levers of democracy. Examples include the use of digital media to: broaden the public's engagement with indigenous people and their lives, give more exposure to women in politics, build well-networked, educated and empowered communities, and encourage political engagement from youth.

The second category of opportunities relate to digital media's use by people in governments to make the processes of democracy more inclusive, to increase engagement with citizens, improve transparency of government work, and rebuild trust in democratic processes. Examples of such work include online deliberative democracy processes, open or e-government initiatives, and funding of public service journalism, platformed on digital media.

DEMOCRATISATION OF INFORMATION PUBLISHING

Digital media enables the creation and sharing of content by anyone. This aspect in particular, the literature shows, has the potential to improve democratic participation by facilitating dialogue both between governments and citizens (improving institutional trust) and between otherwise divergent groups and individuals in society.

BROADENING OF THE PUBLIC SPHERE

The literature suggests that digital media can be used to widen policy conversations to include marginalised individuals and communities who have been previously excluded from democratic processes. A good example of this is the @IndigenousX Twitter account.

INCREASING EQUALITY OF ACCESS TO AND PARTICIPATION WITHIN POLITICAL PROCESSES

Several studies found that digital media increases equality of access to and participation within political processes, in terms of gender, class, race and age. Specifically digital and social media:

- gives more positive exposure to women politicians than traditional news media;
 - builds well-networked, educated and empowered communities, which may previously have been economically and socially marginalised by digital divides, (when incorporated with other good government policies such as civics education),
 - facilitates the formation of both 'ad-hoc' and longer-term, group-based online communities focused on fighting racism, which can provide a safe space of belonging for ethnic minority groups.
 - softens political inequality patterns by encouraging political engagement from 16-29 year olds.
-

INCREASING ENGAGEMENT IN POLITICAL PROCESSES

Numerous studies found links between digital media and increased engagement in political processes in the general population, not only in marginalised groups. This includes engagement in elections, different forms of deliberative democracy, as well as participation in more informal political action such as protests.

INCREASING TRANSPARENCY AND ACCOUNTABILITY FROM GOVERNMENT

While digital media has the potential to help rebuild trust in public trust in democratic institutions and policies using “open government” and “e-government” initiatives, the research suggests it is likely dependent on the base level of trust (i.e such approaches may be more effective in low trust environments where there is less initial transparency).

PROMOTION OF DEMOCRATIC VALUES

People in governments have been able to actively promote democratic values, informed debate, tolerance and respect for other groups using digital media. Examples of direct action include government funding of public service journalism, funding of independent statutory organisations such as All Together Now in Australia, which encourages embracing of cultural diversity. Less direct action includes the use of digital media to promote participatory democracy activities, e.g deliberative forums.

SUMMARY

The opportunities for digital media are significant and important. If used well, digital media can enable governments to respond effectively to the experiences of marginalised groups, to ensure equitable policies and practices are designed, delivered and adjusted, and to build trust in the democratic institution as responding to the needs of all people. It offers as much to people pushing against barriers to their progression, inclusion, and improved wellbeing in society as it does to people in government looking to remove those barriers and build a more inclusive democratic system.

THE THREATS

The threats to this promise outlined in the literature are significant however, and most are intricately bound up with the concentration of power in profit-driven companies. The seven key threats we identified to inclusive democracy from digital media were: the increasing power of private platforms, foreign government interference in democratic processes, surveillance and data protection issues, fake news, misinformation and disinformation, filter bubbles and echo chambers, hate speech and trolling, and distrust/dissatisfaction with democracy.

Some of these threats or problems originate in the structures and systems of society e.g the power of private platforms over people's lives. Others operate at an individual level e.g. a growing distrust of democracy. However, all these threats are interconnected. Together they threaten to derail the democratic promise of digital media.

INCREASING POWER OF PRIVATE PLATFORMS

Private platforms have increasing power to determine all aspects of our access to information, social interactions, and democratic activities. Researchers highlighted the increasing dominance of an increasingly small number of privately-owned platforms over the internet. People who own and control these platforms have a monopolisation tendency linked to the relationship between the mining of user-data and their imperative to make profits. This model of operation is termed “platform monopolies”. The monopolisation tendency makes alternatives to the data-extraction for profit model, for example co-operative, democratised ownership models, hard to start up and survive.

The concentrated power of these platforms shapes not just the wider information context and ability to develop alternative non-extractive models of digital information provision and sharing, but individual’s personal lives also. Platform monopolies affect how and with whom we interact socially through algorithms. A body of literature points to the actions that these people in this companies take that impact human rights, both through the control of personal data and the level of control over what appears in the public sphere.

From this model of platform monopolies flows a series of further threats to democracy. Some relate to the features of the platforms, directly linked to the capture of people’s personal data. The collection and on-sale of personal data by these platforms, to both governments looking to undertake surveillance on their own citizens, and private organisations wishing to make profits, erodes public trust in information systems, and curtails the professional work of the media and writers - a key plank in our democracy.

FOREIGN GOVERNMENT INTERFERENCE IN DEMOCRATIC PROCESSES

The literature shows that the interference in democracy, specifically through the use and manipulation of digital and social media contributes to decreased turnout and voter disengagement. Disinformation campaigns by foreign governments exaggerate already existing tensions and polarisations and encourage a lack of faith in the electoral system and lack of trust in the idea of liberal democracy.

SURVEILLANCE AND DATA CAPTURE

Evidence shows that the provision of private data to both governments for the purposes of surveillance and private organisations for profit-making by platforms has a curtailing effect on key components of democratic function. Specifically, there has been a demonstrated curtailing effect on the private and public practices of both writers and journalists. While surveillance of Muslim communities for example contributes to alienation from mainstream society.

FAKE NEWS, MISINFORMATION AND DISINFORMATION

The creation of the “attention economy” also poses a significant threat. People’s propensity to attend to shocking, false, or emotive information, especially political information, is exploited and used as a commodity product by digital media platforms.

The literature shows that governments with the means and inclination to manipulate information can tailor false information towards individuals with the express intent of interfering in other countries democratic processes, for example Russian government interference in the U.S election of 2016 using ‘bots’ and disinformation campaigns. While misinformation and disinformation, especially political disinformation, targeted at individuals on digital media, is used to influence politics, from national elections through to information provision and sharing with regard to political issues and policy more generally. Political misinformation in particular has been found to have a significant direct and indirect impact on democratic participation and engagement.

FILTER BUBBLES AND ECHO CHAMBERS

Filter bubbles are specific technical effect of the attention economy. Facebook's news feed is a filter bubble, created by a machine-learning algorithm which draws on data created by user networks, likes and comments and how much organisations are willing to pay to be present there.

Filter bubbles follow a longer-term trajectory within advertising (including political advertising and now disinformation) which has sought to collect data in order to tailor adverts to target groups, however, now they can be targeted to specific individuals. This can contribute to the formation of echo chambers, which is the reinforcement of existing beliefs (confirmation bias) through selective exposure to information. Hence, the technical and economic drivers of filter bubbles can act to reinforce echo chambers.

Increasing numbers of automated social media 'bots' have been linked with the spread of political disinformation and thus the reinforcement of echo chambers. Filter bubbles and the related echo chambers they feed into are linked to a decline in trust in the ability to traditional news media to provide reliable information. They have been found to exacerbate political divisions and polarisation, and have negative implications for the mechanisms of liberal democracy. Developing a broad consensus around decisions made in the public good becomes increasingly difficult.

HATE SPEECH AND TROLLING

The rise of hate speech and trolling is linked to polarising effects of filter bubbles and echo chambers. A troll is an anonymous user who deliberately provokes antagonistic reactions for sheer enjoyment. Trolling is aided both by the ease of creating anonymous online profiles and by the atomised nature of internet interaction. Trolling can pose a direct threat to opportunities when it becomes systematically targeted towards minority groups in order to deliberately cause emotional distress. Remaining anonymous makes individuals more likely to escape prosecution for the more egregious examples

Racialised hate speech (otherwise known as cyber racism) is specifically targeted towards ethnic minority groups, and has become increasingly coordinated in recent years, through the rise of the "alt-right". It encourages affected groups to retreat to safe locations, rather than engaging with national debates and institutions.

Sexualised hate speech is primarily targeted towards women (together with members of the LGBTIQ community), and is characterised by its specifically misogynistic nature. It is often directed towards women in the public eye, or those in influential positions, such as journalists, with proponents directing critical attention onto their supposed essential gender characteristics, rather than their work. It has a negative impact on efforts towards the broadening of the public sphere, as women are discouraged from writing what they may feel are controversial stories.

More generally, research has found a correlation between strong, vocal disagreements with an individual's perspectives and a "spiral of silence" which acts to curtail the voicing of contentious opinions by minority groups. The particular ability of trolls and hate speech to fan antagonistic "flames" rather than promote rational debates, has a direct impact on democratic participation.

DISTRUST/ DISSATISFACTION WITH DEMOCRACY

While distrusts with democratic process is a longer-term issue, digital media has likely exacerbated this pattern across western democracies. Researchers argue that trust, informed dialogue, mutual consent, and participation- fundamental features of democracy are being eroded by the features that make social media so profitable. Researchers also found that the way in which the information is distributed on digital media (horizontal, and decentralised and interactive) increases intolerance of others, polarisation and scepticism toward democracy

SUMMARY

The opportunities of digital media, while still apparent, appear to have been suppressed by the sheer weight of fake news, filter bubbles, populism, polarisation, hate speech, trolls and bots, that have emerged from the concentration of power in a small group of private organisations seeking to maximise profits. Digital platforms initially celebrated for their democratic possibilities have transformed into anti-democratic power centres through the collection and exploitation of users attention and data. These privately owned platforms have largely escaped public oversight or regulation over their ability to harness this new power for commercial or political gain.

The question is what can policymakers do to re-calibrate? Are there empirically tested public policies and approaches that can ensure digital media works to strengthen and deepen democracy?

THE SOLUTIONS

A HIERARCHY FOR SOLUTIONS AND INTERVENTIONS

At The Workshop, we take an evidence-informed, hierarchical approach to exploring and understanding problems, and investigating and analysing solutions, policies and practices to overcome them. We work especially to highlight the critical role of structures and systems in improving people's lives with the least individual effort required (though not the least political effort).

IDENTIFYING DRIVERS OF THE PROBLEM

First we ask is the problem we have identified a structural or systems level problem (e.g. the structure of the economic model, the power of private markets over people's wellbeing) or a group/individual level response to the issue (e.g. distrust in democracy that results from a lack of inclusion in democracy)? Sometimes defining where problems originate is complex as there are interactions and feedback loops, as with all complex issues.

For example, hate speech is an individual or group behaviour, it is fundamentally about how people or institutions treat others, however the upstream issues that encourage and enable hate speech, intolerance and bigotry must be explored. Wealth, gender and ethnic inequalities in society, for example, mean digital platforms are primarily owned, designed and managed by those with little experience of differential or harmful treatment based on their position in society.

Without knowledge of how power imbalances and differential treatment based on gender or race play out in society, or a commitment to overcoming them, people who control these platforms can design in policies and practices that encourage hate speech and trolling. By presenting problems in a hierarchy we endeavour to make the feedback loops and upstream structures and systems issues clearer to people.

IDENTIFYING WHERE PEOPLE SHOULD INTERVENE FOR GREATEST IMPACT

In terms of considering "what works", we focus on 'upstream' or structural and systems responses and solutions to the problems. We take this approach because research from across disciplines focussed on enhancing population wellbeing and equity shows interventions at this level:

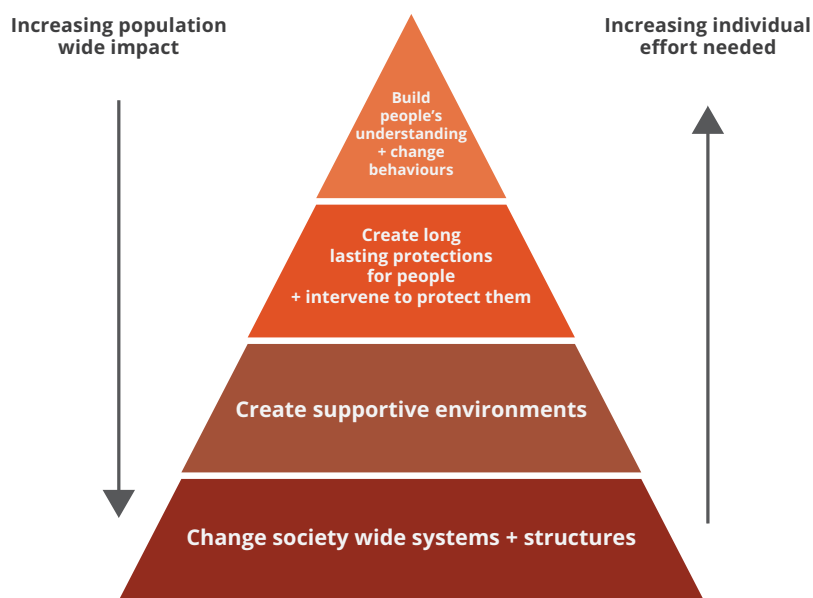
- have the most significant impact on most people's lives and outcomes, and
- require the least effort from individuals to achieve change, and the least resources from those trying to implement change.

We place less emphasis on individual behavioural solutions, not because they are not effective, but because to be effective these solutions (e.g civics education, or consumers closing their Facebook accounts) take significant effort from both individuals and those encouraging such action, and may not address the structural drivers that cause the problems upstream. In addition, people expending energy on individual level solutions can divert energy from investing in understanding and acting on structural level solutions.

The possible interventions identified both in the literature review and by the interviewees are discussed in the context of a hierarchy from those likely to be most effective and requiring least individual effort, to those likely to have the least impact and requiring most individual effort on a population-wide scale. This hierarchy comprises interventions that:

- Change society-wide structural & systems issues to re-establish citizen power
- Create supportive environments & contexts - making the default digital space inclusive and safe
- Create long-lasting protections for people, and intervene to protect them from digital threats
- Build understanding of digital media threats and change individual behaviours in response.

Wellbeing Impact Pyramid



Adapted by The Workshop from Frieden (2010)

FINDINGS

While our literature review was not exhaustive, in general we found a dearth of empirically tested solutions. Likewise, and possibly because of the lack of evidence in the literature, the experts interviewed for this research generally had more to say about the risks and threats they saw arising from digital media than they did on potential workable solutions. However, we did find some common ground between the literature and in the interviews, in terms of solutions.

In line with The Workshop's evidence-led approach set out above, we discuss what empirical evidence we did find in a hierarchy, starting with those solutions likely to have the greatest impact and require least individual effort.

CHANGE SOCIETY-WIDE STRUCTURAL & SYSTEMS ISSUES TO RE-ESTABLISH CITIZEN POWER.

This section focuses on structural and systemic change, addressing for example the disproportionate power of the tech giants vis-a-vis governments, citizens and their domestic competitors.

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Regulating platforms like other industries. Currently, regulatory debates largely centre around defining the structure, terms and conditions of what kind of industry private intermediaries represent. How platforms should be regulated or governed thus partly hinges on how these services are defined; for example, whether social media platforms are media companies, public spaces, utilities or some other service largely informs how they can ultimately be governed. There is little or no empirical evidence to show how regulation in this area would or would not work, and therefore adaptive approaches to policy and regulation will be needed. This will involve ensuring that the impacts of any change are regularly monitored and changes made as needed in response to those findings.

Introducing new modes of collective action. Under industrial capitalism we had collective bargaining, the strike – e.g., forms of collective action that were sanctioned by law and had the support of a society that allowed people to tame capitalism with legal protection. In relation to digital media researchers suggest there are opportunities for more collective action both by tech workers, demanding for example more ethical design in the products they work on, and by digital media users. New forms of collective, collaborative action that connect users/consumers with the market and state to tame and outlaw surveillance capitalism are suggested by multiple researchers, but again there is no empirical testing to yet drawn upon.

COMBAT FAKE NEWS BY:

Supporting a vibrant and diverse media sphere. One that balances strong, independent and adequately resourced public service media with a non-concentrated commercial media sector. Although there is an existing body of research showing the positive impact of a vibrant and healthy public and independent media on democracy, the specific impact of investing in media in the context of digital media is a widely proposed but as yet unmeasured idea.

CREATE SUPPORTIVE ENVIRONMENTS & CONTEXTS - MAKING THE DEFAULT DIGITAL SPACE INCLUSIVE AND SAFE

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Designing new competitive digital media solutions. Disruptive technology is needed to forge an alternative digital future that in turn, facilitates a more democratic internet. This means the creation of platforms offering a different set of affordances (ie not those driven by platform monopolies). Platform cooperatives like Loomio subscription-based models and pro-privacy and non-commercial alternatives are already in use and show some evidence of effectiveness in the literature.

REDUCE INTERFERENCE FROM FOREIGN GOVERNMENTS AND POWERS BY:

Designing new anti-cybersecurity infrastructure and drawing upon “big datasets” to review and assess electoral policies. The research in this area is also largely normative, and seems to generally prescribe such infrastructure and reviews will reduce threats to elections and other political processes.

ADDRESS SURVEILLANCE & IMPROVE DATA PROTECTION BY:

Regulating companies’ information management practices. Some regulatory measures, like the Singaporean Data Protection Act 2012, work to and have been proven effective in bringing formal charges to data mismanagement and abuse.

Making regulatory changes to data privacy policies. However, there is little evidence to suggest that these changes will reduce surveillance/data collection so much as regulate how that data is stored, accessed and used by data collectors and other third parties

COMBAT FAKE NEWS BY:

Developing and circulating persuasive counter-narratives. The focus would need to be on emotional not rational, appeal. This is proposed but unmeasured.

OVERCOME FILTER BUBBLES/ECHO CHAMBERS BY:

Supporting new platform designs with different design affordances

Design affordances ascribe meaning to how to use the digital media tool, for example Facebook has a “friend” button directing the user towards ways of interacting based on mutual agreement, also a “share” button, while Twitter has a “follow” button, open to all people using the platform, directing or suggesting different ways of interacting. The design of these affordances has an impact on inclusion and participation, as well as the types of interactions people experience and information they are exposed to. There is some suggestion that design affordances can reduce the effects of filter bubbles by engaging internet users in more ideologically diverse communities.

Non-commercial platforms like Loomio, for example, afford different modes of interaction based on the features (e.g., tools, interface) and environment (e.g., deliberative; asynchronous) it makes available outside a commercial space. The platform affords an environment of less performative, and thus more considered, dialogue, discussion and

debate. The relationship between design and civility on these new platforms has been empirically demonstrated, shown a reduced propensity to engage with similar-minded people encouraged by automated filter bubbles and to move deliberation beyond debate to collective agreement.

OVERCOMING SILENCING EFFECTS OF HATE SPEECH BY:

Supporting new platform designs with different design affordances. Well-designed, collectively-owned, online deliberative fora like Loomio have been empirically shown to also create a safe environment for marginalised groups.

Research suggests that intentionally building more participatory forms of engagement into platforms might reduce filter bubbles, echo chambers and incivility (particularly on mobile devices), while increasing communication and deliberative processes. Therefore, the act of consciously designing social platforms to engender pro-social forms of engagement can have a demonstrated impact on civility

IMPROVE TRUST IN DEMOCRACY BY:

The creation, selection and use of online platforms that afford citizen participation and deliberation. Some empirical evidence shows that direct and participatory democratic engagement/processes, e-government, and open government improve trusts. International research has found that engaging citizens in deliberative processes often results in profound changes in deliberating citizens' "frequently in the direction of more common good-oriented policies", but for them to be effective the systems and platforms used in these deliberative processes must also enable these practices to emerge. The techno-social affordances inherent to different online platforms affect and shape the nature of engagement, deliberation and discussions.

Using digital government processes. Transparent, easy to access and well designed e-government and open government initiatives have been shown to increase positive feelings and citizen trust in local government. Some evidence shows governments that have created usable, intelligible websites, and offer non-exclusionary solutions for those lacking computer and internet access or basic digital literacy skills, have been most successful in their e-government initiatives and constituent satisfaction.

CREATE LONG-LASTING PROTECTIONS FOR PEOPLE OR INTERVENE TO PROTECT THEM FROM DIGITAL THREATS

REDUCE THE POWER OF PRIVATE PLATFORMS, COMBAT INCIVILITY AND MISINFORMATION ONLINE BY:

Improving Content Moderation. Calls for new regulatory policies around content moderation at large intermediaries acknowledge content moderation remains an opaque and difficult practice, and on its own is not a fix-all solution. Current policies at the largest intermediaries attempt to balance stakeholder expectations (including users, consumers, advertisers, shareholders, the general public), commercial business goals, and jurisdictional norms and legal demands (which are generally governed by liberal-democratic (US) notions of “free speech”), goals related to inclusive and participatory democracy are not included.

The most common ‘workable solution’ presented as it relates to content moderation are processes that combine technical and social (human) responses. However, advances in semi- or fully automated systems, including deep learning, show increased promise in identifying inappropriate content and drastically reducing the number of messages human moderators then need to review. In the literature however, researchers note that neither automated nor manual classifications systems can ever be “neutral” or free from human bias. Human and/or automated content moderation is unlikely to achieve “civil discourse,” a “sanitised” internet or other speech and engagement goals through moderation alone. Therefore, the combination of automated classification and deletion systems and human efforts remains the most effective content moderation strategy currently on offer. In the few places where they exist government regulations on private intermediaries’ moderation practices have not been empirically tested for their efficacy or effectiveness.

COMBAT FAKE NEWS BY:

A multi-stakeholder content moderation. This is an approach that combines human and technical intervention, however this is a proposed but untested solution.

REDUCE HATE SPEECH/TROLLING BY:

Using identity verification systems. Sites that do not allow anonymisation and force pre-registration have been shown to solicit qualitatively better, but quantitatively fewer, user comments because of the extra effort required for engaging in discussion. Empirical research has also found that abusive comments are minimised when anonymous commenting is prohibited.

BUILD UNDERSTANDING OF DIGITAL MEDIA THREATS AND CHANGE INDIVIDUAL BEHAVIOURS IN RESPONSE.

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Building citizen-consumer activism and creating a “sea change in public opinion”. Scholars and theorists suggest that a shift in public attitudes is needed to persuade digital media companies to change, there is however no empirical data to draw upon as to how effective this approach would or would not be.

ADDRESS SURVEILLANCE AND DATA PRIVACY ISSUES BY:

Encouraging individuals to employ technical solutions. Such solutions include ad-blockers and ad-tracking browser extensions, private browser options (e.g. Tor), open source platforms and cooperative platform models. “Evidence” supporting the efficacy of these tools and alternatives, however, is typically anecdotal or prescriptive in nature (as opposed to empirical).

COMBAT FAKE NEWS BY:

Education, particularly around critical thinking. Evidence has emerged in health for this approach.

REDUCE HATE SPEECH BY:

Building Resilience through Support Networks. Developing fast and effective reporting mechanisms and support networks, e.g Advocacy and civil society organisations like All Together Now, have demonstrated some success with building online reporting tools that rely on crowdsourcing to identify – in order to remove - racist hate speech online. A networked approach can effectively combat the effects of hate speech; by building counter-narratives that counteract racism for example.

Coordinating diverse stakeholders to apply pressure to private intermediaries, in ‘long-haul’ campaigns, has also been effective in having hateful content removed from social media. Speed of removal is considered essential to diffusing the power of hate speech and trolling. Pressure from researchers and advocacy groups alike have also encouraged some platforms to design more pro-social tools (i.e., affordances) into their systems.

IMPROVE TRUST IN DEMOCRACY BY:

Civics education. Educating children in schools on “good citizenship” has been positively associated with increased political engagement.

WHAT ROLE FOR NZ?

All of this raises the question: what role should New Zealand play in the wider global efforts to respond to the challenges of digital media? Some interviewees argue that New Zealand should follow the lead of bigger, like-minded liberal democracies like the United Kingdom, the European Union and Australia. Others thought New Zealand could, and should be leading on these issues. Some saw specific opportunities for New Zealand to provide leadership in niches, like indigenous data sovereignty.

One thing many people agreed on was a sense of urgency – an urgency which has increased considerably in the months since many of these interviews took place.

As one participant put it, “we’ve got some really resounding early warning signals about how this stuff can be used to erode our democratic institutions, and if we don’t sit up and take notice of it, and don’t provide the necessary technical, social, and regulatory responses, we might wake up and find that we’ve missed the opportunity.”

NEW ZEALAND AS FOLLOWER

One of the common reasons given for taking the path of following the lead of others was New Zealand's size. However another, perhaps more critical, argument was that New Zealand would need a much better system for making policy on these issues before we can be any kind of global leader. Before we can lead, this participant argued, we need to build up our national capacity to understand and deal with these issues, and build up more of an evidence base, we need to be equipped to have a solid base for developing policy ourselves. It's hoped this research can help contribute to that process.

NEW ZEALAND AS LEADER

On the other hand, some interviewees asked why New Zealand should be a 'taker' of policy on these issues, and identified a great opportunity for New Zealand to team up with other like-minded democracies. We are typically at the cutting edge of technology, they argued, so why not take a lead on this. Digital media has brought advantages to New Zealand, they argued, so we want to make sure that we don't lose the upsides of the new digital economy. Playing a leading role in the global response to the threats of digital media can help ensure that we do not.

NEW ZEALAND AS NICHE INFLUENCER

Some interviewees pointed to New Zealand's track record of taking a principled stand on big global issues, giving our nuclear-free policy as an example. One example given as an area in which New Zealand could show leadership is in the development of a tech workers' union. Because New Zealand has comparatively better employment protections than many other places where tech people work, they said, we already have less of the fear of speaking up. We also have a small enough sector where personal relationships can very easily be brought to bear on these situations.

Finally, but perhaps most importantly, some interviewees argued that there was a role for New Zealand to play as a leader on indigenous data sovereignty and issues relating to Māori digital issues. This would first require us to address the significant gaps in our own protection of indigenous rights online. One of the most critical issues is the need to protect indigenous data sovereignty, allowing Māori ownership and control of Māori data.

CONCLUSIONS

At the heart of the challenges to democracy posed by digital media are three core problems:

1. **Platform monopolies:** two or three corporations control not only our means of communication, but also the content which is distributed both of which are core aspects of our democracy, whilst the market power and global mobility of these companies make it possible for them to avoid national regulatory measures either by moving operations elsewhere or simply ignoring them;
2. **Algorithmic opacity:** algorithmic engines are using huge quantities of personal data to make ever more precise predictions about what we want to see and hear, and having ever increasing influence over what we think and do, with little transparency about how they work or accountability for their impact; and
3. **Attention economy:** the dominant business model of digital media prioritises the amplification of whatever content is best at grabbing our attention, while avoiding responsibility for the impact that content has on our collective wellbeing and our democracy. And the negative impact is brutally clear from both the literature and the world around us.

Combined, these problems pose serious threats to our democracy, so it's critical that our responses to them don't further undermine our democratic institutions. The history of digital media has shown that good intentions can, if not informed by the diverse experiences of users and the research evidence, cause more harm.

INCLUSIVE AND TRANSPARENT PROCESSES ARE CRITICAL

Firstly, as Natasha Tusikov and Blayne Haggart have argued, decisions about what kinds of information we have access to should not be made by a handful of American companies. Nor should our government's role in those decisions take place in backroom negotiations. We need to use democratic processes, which provide some degree of transparency about the decisions being made, accountability as to their impacts, and opportunities for challenge and judicial review. These processes must include meaningful participation by diverse representatives of the people whose lives are impacted by digital media.

MORE RESEARCH NEEDED

Secondly, the stakes are high here, so we must draw on the evidence as to what is most likely to work, where it exists. Perhaps the most predictable finding of this research is that there has been little or no investment by people in government or other research funders into experimenting with and recording possible solutions, and there needs to be more. A list of possible areas for further research is included in the full report.

It's not surprising that there is so little experimental evidence as to the effectiveness of various solutions proposed in the normative literature. Change happens very quickly in this area. Until very recently there has been little investment in research from governments, which would be one of the expected sources of funding for such investigations. More is now urgently needed.

EVIDENCE-LED AND PRINCIPLED APPROACH TO URGENT POLICY DEVELOPMENT

Even in the absence of specific evidence as to the effectiveness of different interventions, there are areas in which action is urgently needed. In those cases, there are key principles that can be followed to reduce the risk of implementing solutions that do more harm than good. As a primary principle, we can take an evidence-informed, hierarchical approach to exploring and understanding problems, and investigating and analysing solutions, policies and practices to overcome them. This involves identifying underlying drivers of the problem, and those interventions which are most likely to have the greatest impact.

We set these principles out in more detail in the section on solutions above, but the key point is that we need to focus on tackling the structural drivers that underlie all the more specific problems outlined above - such as online abuse, the spread of disinformation, radicalisation and polarisation, political interference and manipulation or distraction. Solutions should then be designed to intervene at that structural level addressing and rebalancing power through, for example, governance structures, regulation to restore transparency, accountability and fair competition and genuinely participatory and representative multi-stakeholder processes.

None of this is to say that design solutions and platform affordances are not important. As the research shows, they will be essential. But without some rebalancing of power, without increasing the diversity of people involved in decision-making at the highest levels, those design solutions run the risk of replicating very similar problems to those we now face.

APPLY HUMAN RIGHTS PRINCIPLES

Human rights principles should also be applied to policy development in this area, and are particularly useful where there is an absence of specific research evidence. These principles include:

- > **Universality:** Human rights must be afforded to everyone, without exception.
- > **Indivisibility:** Human rights are indivisible and interdependent, which means in order to guarantee civil and political rights, governments must also ensure economic, social and cultural rights (and vice versa).
- > **Participation:** People have a right to participate in how decisions are made regarding protection of their rights. Governments must engage and support the participation of civil society on these issues.
- > **Accountability:** Governments must create mechanisms of accountability for the enforcement of rights. There must be effective measures put in place for accountability if those standards are not met.
- > **Transparency:** Transparency means governments must be open about all information and decision-making processes related to rights. People must be able to understand how major decisions affecting their rights are made and how public institutions responsible for implementing rights are managed.
- > **Non-Discrimination:** Human rights must be guaranteed without discrimination of any kind. This includes not only purposeful discrimination, but also protection from policies and practices which may have a discriminatory effect.

Each of these principles should be applied in the development of a multi-stakeholder response to the threats to democracy posed by digital media.

AGILE AND RESPONSIVE APPROACH TO POLICY

Finally, in the absence of a strong evidence base, it makes sense to take an agile, iterative approach to policy change. Experiment with all the policies all the time. Ensure that the funding, design, and implementation of policies reflect a record, learn, and adapt approach to measure the impact of any new initiatives or regulations, and to make adjustments as evidence becomes available as to impact.

URGENT AREAS FOR ACTION

Some of the areas in which action is needed include efforts to:

- > **Restore** a genuinely multi-stakeholder approach to internet governance, including rebalancing power through meaningful mechanisms for collective engagement by citizens/users;
- > **Refresh** antitrust & competition regulation, taxation regimes and related enforcement mechanisms to align them across like-minded liberal democracies and restore competitive fairness, with a particular focus on public interest media;
- > **Recommit** to publicly funded democratic infrastructure including public interest media and the creation, selection and use of online platforms that afford citizen participation and deliberation.;
- > **Regulate** for greater transparency and accountability from the platforms including algorithmic transparency and great accountability for verifying the sources of political advertising;
- > **Revisit** regulation of privacy and data protection to better protect indigenous rights to data sovereignty and redress the failures of a consent-based approach to data management; and
- > **Recalibrate** policies and protections to address not only individual rights and privacy but also to collective impact wellbeing. Policies designed to protect people online need to have indigenous thinking at their centre and should also ensure that all public agencies responsible for protecting democracy and human rights online reflect, in their leadership and approaches, the increasing diversity of our country.

In the wake of the Christchurch mosque attacks, a new global momentum has emerged around the role that social media has played in the spread of violent extremism and terrorism, and what can be done to stop it. The New Zealand government has, rightly, stepped up to play a leadership role in that work.

What we need right now is a clear analysis of the wider structural and systemic issues that underpin the immediate moderation challenge and a solid proposal of regulatory and other changes that are needed to tackle those bigger issues. That means ensuring that the current, heightened public debate on the role of digital media in fostering and spreading hate is placed into a wider context of the regulatory and structural changes needed to revive and restore the key features of a healthy and peaceful democracy in our country, and around the world.

Our intention is that this research will help frame, inform and support those efforts.

SUMMARY OF KEY FINDINGS FROM EACH PART OF THE RESEARCH

KEY FINDINGS FROM THE SURVEY

Some key findings from the analysis of the survey data:

- > Use of social media is high, and Facebook dominates with Twitter and Instagram also used.
- > Around a quarter of the sample used social media to engage with ‘political issues’ or politicians.
- > Social media platforms are used for political activity by minority ethnic groups more than Pākehā. Indicating their potential use as a tool for engagement in formal democratic system.
- > Stated trust in news online may be low, but perceptions of information credibility are driven by trust in friends, family and organisations.
- > Most people still rely on mainstream media for information about a key political issue (decriminalisation of marijuana) but friends and family and online news feature strongly.

New Zealanders are relatively accurately informed about the opinions of others with regard to the decriminalisation of marijuana.

There is evidence that New Zealanders who believe their views are in the minority on decriminalisation of marijuana are less willing to share their views both offline and online. This suggests social media platforms replicate rather than overcome existing barriers to engaging in less formal processes of democracy (public discussion and discourse) for people who hold minority views.

KEY FINDINGS FROM THE LITERATURE REVIEW

To paraphrase classical historian Mary Beard, western democracy is a 2000 year old experiment. In 2019 the significant technical disruption that is digital media is having a powerful effect on the results. Yet what is the nature of that effect? Does our collective written and published knowledge tell us what benefits and opportunities digital media offers in building a stronger, more inclusive and participatory democracy? And the threats or risks it poses to it? And what if anything does the empirical evidence tell us optimises the opportunities and reduces the risks to our democracy from digital media?

The answer to these questions remains elusive. While our literature review was not exhaustive, this research confirms that there is, at present, a troubling dearth of scientific, empirical, evidence-based research that tests or aims to validate “workable solutions” to the seven key threats to democracy we’ve identified in this project.

While some empirical evidence exists, notably in the area of designing new platforms and affordances with pro-social intent, the significant majority of the research relating to the threats we identified is based on expert opinion and normative approaches. Meaning, it presents theoretically sound arguments about the way things “ought to be” if democracy is to be “reclaimed” from incivility and a rogue form of capitalism in the digital age.

In the expert opinion literature the following four themes were identified:

1. **Policy / Legal Solutions**

For example, adapt existing legislation; create new legislation; institute new oversight bodies or inter-government agencies; or to improve regulations on content moderation.

2. **More Corporate Transparency**

Currently the lack of transparency around moderation practices presents challenges to accountability, governance, and the ability to apply public and legal pressure. Expanding empirical research to improve moderation processes requires private intermediaries to make these processes and practices accessible to researchers.

3. **Better Design**

Platform design can influence the way individuals, organisations and institutions make decisions around platform uses/objectives. Pro-social and democratic values must be encoded into the infrastructure of the internet, including algorithms. At present, the normative values embedded into these global private intermediaries – e.g., openness, connectedness, free speech, etc. – are not culture-neutral norms. It the first step towards designing more deliberative spaces, pro-social tools and online environments.

4. **Improve Content Moderation**

Calls range from the standardisation of industry-wide “best practices” to more transparency and researcher access. These actions would require greater corporate transparency, corporate grievance mechanisms that are transparent, accessible and in accordance with international human rights law, and multi-stakeholder, and inclusive governance approach, and content moderation should become an organisational priority rather than department silo.

This absence of tested solutions is not evidence that proposals do not work, but that they are untested. This leads us to conclude there is a critical need for investment in more research. People in government, civil society, NGOs, and private enterprise need to commit to researchers and projects who will do pre-and post-testing of solutions that stakeholders are recommending.

Such research will not only measure effect and enable us to extend what's working to other places or contexts, but ensure future normative prescriptions are informed by evidence beyond the anecdotal (or budgeting restrictions).

It is critical that people in the New Zealand government especially measure whether or not what is being done is working to build a more inclusive and participatory democracy. New Zealand would break significant ground in that regard.

When people in government and civil society seek recommendations for solutions, they need to mitigate the risk that experts reproduce "solutions" that fit the professional discourses in which they're embedded. To do this, it is important that people in government ask multi-stakeholder group participants:

- > What if any evidence do they have for the suggestions made?
- > What experiences inform these recommendations and why do they identify them as workable solutions over others?
- > How do they imagine testing their effectiveness?
- > Given the current lack of evidence, it is critical that the values, experiences, and outcomes that underlie recommendations are made transparent and visible.

KEY FINDINGS FROM THE INTERVIEWS

One of the questions we posed in the interviews was what had changed in the landscape of democracy, through the influence of digital media, and what has not changed.

A FAMILIAR BUT CHANGING LANDSCAPE

Many things have not changed, participants told us. Misinformation, disinformation and harassment are not new. Outrage, political polarisation and extremism are not new. Filter bubbles, soundbite politics and data capture - none of this is new. Even the erosion of the authority of published material isn't new. And perhaps most tellingly, the cultural hegemony of tech isn't new. Some participants argued that despite all that has changed as a result of digital media, the replication of existing power structures in the governance and management of the tech giants has inevitably lead to reinforcing many already entrenched power imbalances. Further, they said, the lack of diversity at the governance and senior management level prevented these companies from identifying and responded adequately to the risks and threats inherent in their platform designs.

So what has changed? While recognising that the foundations of mis- and disinformation, online harassment and abuse, polarisation and extremism all existed well before the rise of digital media, most interviewees nonetheless saw particular ways in which the

features and functions of digital media has changed the scale, intensity and reach of those phenomena. Digital media has changed the scale, speed and breadth at which information can be shared. It has allowed advertisers, including political advertisers, to target people with much greater precision. Digital media has generated new levels of distraction, undermining citizens capacity to engage in the complex thinking demanded in a democracy. Data has taken on a new value, and has been gathered and used at an unprecedented scale. And finally, but again, perhaps most importantly, a very small number of very large companies control the means of communication used by the majority of people in most democracies on the planet.

So given what has changed with the rise and digital media, what has stayed the same, and the structural underpinnings of the major digital platforms - where are the biggest opportunities for democracy? The obvious and most commonly cited opportunities were in the democratisation of information, increased diversity in public discourse, more public engagement with government and democratic process, and in increased transparency and openness in government.

On the other hand, participants described considerable risks and threats to democracy including digital exclusion. The most commonly cited risks were the impact of digital monopolies, lack of competition and their impact on public-interest media and misinformation and disinformation, including deepfakes and the consequent erosion of trust in information. Other commonly cited risks include political manipulation including foreign interference, cybersecurity of government and security of elections, and the more common manipulation through political advertising, and related risks of polarisation, radicalisation and 'echo chambers'. Other significant risks highlighted by participants were the impact on democracy of online abuse and hate, disengagement, distractions and attention hijacking, and loss of privacy and consent fatigue. Woven throughout many interviews was a recognition that a lack of transparency and accountability by the big platform companies underpinned and exacerbated all of these risks.

As one participant put it, overall, the picture of how democracy as a form is evolving under the influence of digital media is 'quite messy'. "[I]t's got all these new ways to participate, all these new channels for participation. At the same time, it's getting harder to curate and access that content online, and also critique it. So it's a messy space to talk about risks and opportunities, because the whole landscape is so complicated and moving."

SOLUTIONS

Interviewees suggested a range of interventions and solutions to both maximise the opportunities for democracy presented by digital media and minimise the threats. These range from interventions at the structural and systemic level through to suggestions for individual behavioural change.

Some of the areas in which action was identified as being most urgent include effort to:

- **Restore** a genuinely multi-stakeholder approach to internet governance, including rebalancing power through meaningful mechanisms for collective engagement by citizens/users;
- **Refresh** antitrust & competition regulation, taxation regimes and related enforcement mechanisms to align them across like-minded liberal democracies and restore competitive fairness, with a particular focus on public interest media;
- **Recommit** to publicly funded democratic infrastructure including public interest media and the creation, selection and use of online platforms that afford citizen participation and deliberation.;
- **Regulate** for greater transparency and accountability from the platforms including algorithmic transparency and great accountability for verifying the sources of political advertising;
- **Revisit** regulation of privacy and data protection to better protect indigenous rights to data sovereignty and redress the failures of a consent-based approach to data management; and
- **Recalibrate** policies and protections to address not only individual rights and privacy but also to collective dynamics and wellbeing, and protect indigenous rights. Public agencies responsible for protecting democracy and human rights online should reflect, in their leadership and approaches, the increasing diversity of our country.

